Tools and approaches for mass cytometry data analysis

Erin Simonds, PhD
W. A. Weiss Lab, UCSF

May 23, 2013
Biaxial plots are not a scalable solution

Parameters: 32
Plots: 496
How can we explore in 30+ dimensions?
The cytometrist’s toolbox

Transform raw data
- Histogram
- Biaxial plot
- 3D plot
- Radar

Reduce dimensionality
- PCA
- Gemstone
- SPADE
- viSNE
- Wanderlust

Summarize statistics
- Box plot
- Heatmap

Identify clusters
- FlowClust
- FlowMerge
- FLOCK
- FLAME
- SamSPECTRAL
Vetting the CyTOF: In a fair fight, nearly identical data

**Experimental design**

7 Surface Markers (CD 3, 4, 8, 20, 33, 45RA, 56)
2 Functional Markers (pSTAT3 & 5)

The cytometrist’s toolbox

Transform raw data
- Histogram
- Biaxial plot
- 3D plot
- Radar

Reduce dimensionality
- PCA
- Gemstone
- viSNE
- Wanderlust

Summarize statistics
- Box plot
- Heatmap

Identify clusters
- FlowClust
- FlowMerge
- FLOCK
- FLAME
- SamSPECTRAL
SPADE: Spanning-tree Progression Analysis of Density-normalized Events

Qiu et al., Nature Biotechnology, Oct. 2011
SPADE projects bone marrow as a continuum of phenotypes

CD123

Rediscovery of canonical signaling pathways

Dasatinib differentially affects immune cell subsets

IL7 (pSTAT5)  

PVO₄ (pSTAT5)  

+ Dasatinib  
Abl / Src-family kinase inhibitor

The cytometrist’s toolbox

Transform raw data
- Histogram
- Biaxial plot
- 3D plot
- Radar

Reduce dimensionality
- PCA
- Gemstone
- SPADE
- viSNE
- Wanderlust

Summarize statistics
- Box plot
- Heatmap

Identify clusters
- FlowClust
- FlowMerge
- FLOCK
- FLAME
- SamSPECTRAL
viSNE preserves high-dimensional, non-linear relationships
viSNE maps healthy marrow as discrete populations

Amir et al., Nature Biotechnol. ePub 19 May 2013
viSNE can help detect rare subpopulations

Barcode A: Healthy bone marrow cells
Barcode B: ALL cells spiked into sample at 0.25% frequency (1/400 cells)
Same data, different dimensionality reduction algorithms

Amir et al., *Nature Biotechnol.* ePub 19 May 2013
viSNE is still stochastic, but much more reproducible
SPADE’s stochasticity causes difference between runs

So why should I ever use SPADE?
→ Clustering allows *comparison* between samples

Amir et al., *Nature Biotechnol.* ePub 19 May 2013
The cytometrist’s toolbox

Transform raw data
- Histogram
- Biaxial plot
- 3D plot
- Radar

Reduce dimensionality
- PCA
- Gemstone
- SPADE
- viSNE
- Wanderlust

Summarize statistics
- Box plot
- Heatmap

Identify clusters
- FlowClust
- FlowMerge
- FLOCK
- FLAME
- SamSPECTRAL
Goal: Characterize the GCSF-responsive subpopulation

Who are these cells?
Are they the same in every patient?
Surface marker profiling of G-CSF responders in AML

- **Differential gene expression**
  - Non-responders
  - Responders

- **Markers from literature**
  - CD14
  - CD32
  - CD36
  - CD86
  - CD93
  - CD11a
  - CD1d
  - HLA-DR

- **CD46**
- **CD69**
- **CD117**
- **CD133**
- **CD155**
- **CD321**

- **LSCs**
  - CD13
  - CD47
  - CD15
  - CD64
  - CD33
  - CD114
  - CD34
  - CD123
  - CD38
  - CD11b
  - CD44
  - CXCR4
  - CD45
  - TIM3

- **Screen pediatric AML samples**
  - (18 diagnosis, 3 relapse, 7 healthy)

- **Gate G-CSF responders**

- **Identify enriched markers**
Leukemic GCSF responders have distinct surface profiles

with Rob Bruggner & Sean Bendall
The cytometrist’s toolbox

Transform raw data
- Histogram
- Biaxial plot
- 3D plot
- Radar

Reduce dimensionality
- PCA
- Gemstone
- SPADE
- viSNE
- Wanderlust

Summarize statistics
- Box plot
- Heatmap

Identify clusters
- FlowClust
- FlowMerge
- FLOCK
- FLAME
- SamSPECTRAL
Human B cell development is less understood than in the mouse.

Surface Markers:
- CD34
- CD38
- CD10
- CD10
- IL7R
- CD24
- CD19

Intracellular Markers:
- CD179a
- Rag
- TdT

IgH Genes:
- DJ$_H$
- VDJ$_H$

Adapted – Cobaleda, BioEssays 2009
The predicted trajectory ‘rediscover’ human B cell development

“Wanderlust” algorithm: El-ad Amir, Dana Pe’er (Columbia)
Following the predicted B cell trajectory

- Fluorescent panel design $\rightarrow$ Sorting $\rightarrow$ qPCR

- VDJ rearrangement confirms the trajectory:
  (unrearranged) $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5$ (rearranged)
New insights and resolution in human B cell development

**Surface Markers**
- CD34
- CD38
- CD10
- IL7R
- CD24
- CD19
- Ki67
- CD179b
- CD179a
- Rag
- TdT
- pSTAT5

**Cytoplasmic Markers**

**IgH Genes**
- DJ_H
- VDJ_H
Prof. Garry P. Nolan

Sean Bendall
Kara Davis
Harris Fienberg
Astraea Jager
Rob Bruggner
Rachel Finck

Bernd Bodenmiller (→ U. Zurich)
Eli Zunder
Greg Behbehani

Wendy Fantl

and the entire Nolan lab

Stanford
stanford.edu/group/nolan

Columbia
El-Ad Amir
Jacob Levine
Dana Pe’er

St. Jude Children’s
Amanda Gedman
Ina Radtke
James Downing

Mount Sinai
Michael Linderman

Stanford
Peng Qiu (→ MD Anderson)
Sylvia Plevritis